JOURNAL OF APPROXIMATION THEORY 7, 256–264 (1973)

Best Approximate Solutions on Finite Point Sets of Nonlinear Differential Equations

MYRON S. HENRY

Department of Mathematics, Montana State University, Bozeman, Montana 59715

Communicated by E. W. Cheney

Received November 27, 1970

1. INTRODUCTION

In a recent paper [3] the author considers best approximating on $I = [0, c]$ the unique solution $y(x)$ to

$$Ly \equiv y'' + F(x, y, y') + G(x, y, y') - h(x) = 0 \tag{1}$$

with initial conditions

$$y(0) = \beta_0, \qquad y'(0) = \beta_1. \tag{2}$$

The solution is best approximated in the following sense: if $\mathbf{P}_k = \{P(x, A)\}$, where $A = (\beta_0, \beta_1, a_2, a_3, \ldots, a_k)$, and where

$$P(x, A) = \beta_0 + \beta_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_k x^k,$$

then

$$\| L[y(x)] - L[P(x, A)] \|_I = \sup_I | L[P(x, A)] | \tag{3}$$

is minimized over $\mathbf{P}_k$. That is, (3) is minimized over all polynomials of degree k that satisfy (2). In [3] it is shown that if the operator L in (1) satisfies certain conditions, then the following statements are valid:

(A) There exists a polynomial $P_k(x, A^*) \in \mathbf{P}_k$ such that

$$\| L[P_k(x, A^*)] \|_I = \inf_{\mathbf{P}_k} \sup_I | L[P(x, A)] |. \tag{4}$$

(B) The sequences $\{P_k(x, A^*)\}$ and $\{P_k'(x, A^*)\}$, $k = 1, 2, \ldots$, converge uniformly on I to $y(x)$ and $y'(x)$ respectively.

The choice of $\mathbf{P}_k$ as the minimizing set is not arbitrary. Requiring that the k-th degree polynomials over which (4) is minimized satisfy (2) insures that (B) holds.

Generally it is not possible to find a best approximation from $\mathbf{P}_k$ to $y(x)$ on I (in the sense of (4)), even though one exists. Consequently in this paper we consider best approximating $y(x)$ on arbitrary subsets of I. If $R \subseteq I$, and if the operator L satisfies certain conditions, then it is shown that there exists a $P(x, A_R) \in \mathbf{P}_k$ such that

$$\| L[P(x, A_R)]\|_R = \inf_{\mathbf{P}_k} \sup_R | L[P(x, A)]|. \tag{5}$$

Again the choice of $\mathbf{P}_k$ as the minimizing set is motivated by the desirability of having (B) hold. It is also shown that as the number of points in a finite subset increases, a subsequence of the best approximations to $y(x)$ on these finite point sets converges uniformly on I to a best approximation to $y(x)$ on I.

2. The Operator L

We assume that L satisfies the conditions listed below.

(i) The functions F and G are elements of $C[I \times R^2]$.

(ii) If $P(x, A) = \beta_0 + \beta_1 x + a_2 x^2 + \cdots + a_k x^k$, $A = (\beta_0, \beta_1, a_2, \ldots, a_k)$, and if $\| A \|^2 = \beta_0{}^2 + \beta_1{}^2 + a_2{}^2 + \cdots + a_k{}^2$, then

$$| F(x, P(x, A), P'(x, A))| = 0(\| A \|^\eta) \qquad \text{for large } \| A \|.$$

(iii) There exist functions $u \in C[I]$ and $\varnothing \in C[R^2]$ such that $u(x) \not\equiv 0$ and $\varnothing(y, y') = 0$ iff y or $y' = 0$, and there exists an $\alpha > \max(1, \eta)$ such that

$$| G(x, y, y')| \geqslant r^\alpha \, | u(x) \, \varnothing(y/r, y'/r)| \qquad \text{for all} \quad r \geqslant 1.$$

(iv) The function $h \in C[I]$.

It should be noted that these conditions are essentially those given in [3], and that examples of nonlinear operators L satisfying conditions (i)–(iv) are numerous (see [2, 3, 4] and the example in this paper).

3. Minimizing Polynomials on Arbitrary Point Sets

In this section we establish the existence of best approximations on arbitrary point sets contained in I.

Let

$$K_1 = \{(c_0, c_1) \mid c_0^2 + c_1^2 \leqslant 1\}$$

and

$$K_2 = \{(c_2, ..., c_k) \mid c_2^2 + c_3^2 + \cdots + c_k^2 = 1\}.$$

DEFINITION 1. Let $C = (c_0, c_1, ..., c_k)$, and let

$$P(x, C) = c_0 + c_1 x + c_2 x^2 + \cdots + c_k x^k.$$

If $R \subseteq I$, then

$$\sup_R | u(x) \varnothing [P(x, C), P'(x, C)]| = G_R(C).$$

LEMMA 1. *Let $R, S \subseteq I$ contain at least $k + 1 + l$ distinct points, and suppose that $u(x)$ has at most l distinct zeros on I. Then*

$$\min_{C \in K_1 \times K_2} G_R(C) = \sigma_R > 0, \qquad \textit{and if} \quad R \subseteq S, \quad \textit{then} \quad \sigma_R \leqslant \sigma_S .$$

Proof. Since $R \subseteq S$, $G_R(C) \leqslant G_S(C)$. Thus $\sigma_R \leqslant \sigma_S$. Now suppose that $\sigma_R = 0$. Then there exists a $C^* \in K_1 \times K_2$ such that

$$\sup_R | u(x) \varnothing [P(x, C^*), P'(x, C^*)]| = 0.$$

Then by (iii) either $P(x, C^*) = 0$ or $P'(x, C^*) = 0$ on at least $k + 1$ points for $C^* \in K_1 \times K_2$. Hence either $P(x, C^*) \equiv 0$ on I or $P'(x, C^*) \equiv 0$, on I, a contradiction to the linear independence of $\{1, x, x^2, ..., x^{k-1}, x^k\}$.

THEOREM 1. *Suppose that the set $R \subseteq I$ contains at least $k + 1 + l$ points. If conditions* (i)–(iv) *are satisfied and if $u(x)$ has at most l distinct zeros on I, then there exists a polynomial $P(x, A^*) \in \mathbf{P}_k$, $A^* = (\beta_0, \beta_1, a_2^*, ..., a_k^*)$, such that*

$$\inf_{\mathbf{P}_k} \sup_R | L[P(x, A)]| = \sup_R | L[P(x, A^*)]|.$$

Proof. There exists a sequence $\{P(x, A^{(n)})\} \subseteq \mathbf{P}_k$,

$$P(x, A^{(n)}) = \beta_0 + \beta_1 x + a_2^{(n)} x^2 + \cdots + a_k^{(n)} x^k,$$

$A^{(n)} = (\beta_0, \beta_1, a_1^{(n)}, ..., a_k^{(n)})$, such that

$$\lim_{n \to \infty} \| L[P(x, A^{(n)})]\|_R = \inf_{\mathbf{P}_k} \sup_R | L[P(x, A)]| = \rho_R .$$

Thus for all $n \geqslant n_0$,

$$\| L[P(x, A^{(n)})]\|_R \leqslant \rho_R + 1.$$

Consequently the triangle inequality implies that

$$\begin{aligned} &| G(x, P(x, A^{(n)}), P'(x, A^{(n)}))| \qquad (6)\\ &\qquad \leqslant \rho_R + 1 + | P''(x, A^{(n)})| + | F(x, P(x, A^{(n)}), P'(x, A^{(n)}))| + | h(x)| \end{aligned}$$

for all x in R. Let

$$r_n^2 = \| A^{(n)} \|^2 - (\beta_0^2 + \beta_1^2) = \sum_{j=2}^{k} [a_j^{(n)}]^2,$$

and suppose that $r_n^2 \geqslant \max(1, \beta_0^2 + \beta_1^2)$. Then

$$C^{(n)} = \frac{A^{(n)}}{r_n} = \left(\frac{\beta_0}{r_n}, \frac{\beta_1}{r_n}, \frac{a_2^{(n)}}{r_n}, \ldots, \frac{a_k^{(n)}}{r_n}\right)$$

is an element of $K_1 \times K_2$, $(n = 1, 2, \ldots)$. Also (ii), (iii), (iv), and (6) imply that

$$\begin{aligned} &r_n^\alpha \,| u(x) \varnothing [P(x, A^{(n)}/r_n), P'(x, A^{(n)}/r_n)]|\\ &\qquad \leqslant M_1 + r_n \,| P''(x, A^{(n)}/r_n)| + 0(\| A^{(n)} \|^\eta), \qquad (7) \end{aligned}$$

where $M_1 = \rho_R + 1 + \max_I | h(x)|$. Hence Lemma 1 and (7) imply that $r_n^\alpha \sigma_R \leqslant M_1 + r_n M_2 + M_3 \| A^{(n)} \|^\eta$, where σ_R, M_2, and M_3 are positive constants. Therefore condition (iii) implies that

$$\begin{aligned} r_n^\gamma \sigma_R &\leqslant \frac{M_1}{r_n^{\alpha-\gamma}} + \frac{M_2}{r_n^{\alpha-1-\gamma}} + \frac{M_3}{r_n^{\alpha-\gamma}} (r_n^2 + \beta_0^2 + \beta_1^2)^{\eta/2}\\ &\leqslant \frac{M_1}{r_n^{\alpha-\gamma}} + \frac{M_2}{r_n^{\alpha-1-\gamma}} + \frac{M_3}{r_n^{\alpha-\gamma}} 2^{\eta/2}(r_n^2)^{\eta/2}\\ &\leqslant \frac{M_1}{r_n^{\alpha-\gamma}} + \frac{M_2}{r_n^{\alpha-1-\gamma}} + \frac{M_3 2^{\eta/2}}{r_n^{\alpha-\gamma-\eta}}, \end{aligned}$$

where $\gamma > 0$ and where $\alpha \geqslant \max[\gamma + 1, \gamma + \eta]$. Since the assumption is that $r_n^2 \geqslant \max(1, \beta_0^2 + \beta_1^2)$, the above inequalities imply that $r_n^\gamma \sigma_R \leqslant M_1 + M_2 + M_3 2^{\eta/2}$. Therefore

$$r_n^\gamma \leqslant (M_1 + M_2 + M_3 2^{\eta/2})/\sigma_R = M,$$

where M is a positive constant independent of n.

For each n either $r_n^2 < \max(1, \beta_0^2 + \beta_1^2)$ or $r_n^2 \geqslant \max(1, \beta_0^2 + \beta_1^2)$. Therefore for all n, $r_n^2 \leqslant \max(1, \beta_0^2 + \beta_1^2, M^{2/\gamma})$, and hence

$$\| A^{(n)} \|^2 \leqslant \max[1 + \beta_0^2 + \beta_1^2, 2(\beta_0^2 + \beta_1^2), \beta_0^2 + \beta_1^2 + M^{2/\gamma}],$$

$(n = 1, 2, \ldots)$. Thus the sequence $\{A^{(n)}\}$ is uniformly bounded and hence a subsequence converges. If $A^* = (\beta_0, \beta_1, a_2^*, \ldots, a_k^*)$ is the limit of this subsequence, then $\| L[P(x, A^*)] \|_R = \rho_R$.

4. Convergence of Minimizing Polynomials

Let $\{S_m\}$ be a collection of finite subsets of I, and suppose that

(h_1) $S_m \subseteq S_{m+1}$

(h_2) If $S = \bigcup_{m=1}^{\infty} S_m$, then $\bar{S} = I$.

Set

$$\rho_m = \inf_{\mathbf{P}_k} \sup_{S_m} | L[P(x, A)]| \tag{8}$$

and

$$\rho = \inf_{\mathbf{P}_k} \sup_{I} | L[P(x, A)]|. \tag{9}$$

Because of (h_1, h_2) we may assume without loss of generality that each S_m in the above collection contains at least $k + 1 + l$ distinct points. Then for each m Theorem 1 implies that there exists a $P(x, A_m) \in \mathbf{P}_k$, $A_m = (\beta_0, \beta_1, a_{1m}, a_{2m}, \ldots, a_{km})$ such that

$$\rho_m = \sup_{S_m} | L[P(x, A_m)]|; \tag{10}$$

that is, $P(x, A_m)$ is a best approximation to $y(x)$ on S_m from $\mathbf{P}_k$.

Lemma 2. *Let $\{S_m\}$ be a collection of subsets on I satisfying the hypotheses (h_1, h_2), and let $\{P(x, A_m)\}$, $m = 1, 2, \ldots$, be a sequence of polynomials satisfying* (10) *for each m. Then the sequence $\{A_m\}$, $m = 1, 2, \ldots$, is a uniformly bounded sequence in R^{k+1}.*

Proof. By the reasoning of Theorem 1 we have that if $r_m^2 \geqslant \max(1, \beta_0^2 + \beta_1^2)$, then

$$r_m^\gamma \sigma_m \leqslant M',$$

where $r_m^2 = \| A_m \|^2 - (\beta_0^2 + \beta_1^2)$, $\gamma > 0$, M' is a constant independent m, and σ_m is the positive constant in Lemma 1 with $R = S_m$, $(m = 1, 2, \ldots)$.

Since $S_1 \subseteq S_m$ Lemma 1 implies that $r_m{}^\gamma \sigma_1 \leqslant M'$, and consequently $r_m{}^\gamma \leqslant M'/\sigma_1 = M''$. Therefore $r_m{}^2 \leqslant \max[1, \beta_0{}^2 + \beta_1{}^2, (M'')^{2/\gamma}]$, and consequently

$$\| A_m \|^2 \leqslant \max[1 + \beta_0{}^2 + \beta_1{}^2, 2(\beta_0{}^2 + \beta_1{}^2), \beta_0{}^2 + \beta_1{}^2 + (M'')^{2/\gamma}].$$

That is, $\{A_m\}$ is a uniformly bounded sequence in R^{k+1}.

THEOREM 2. *Let the sequence of sets* $\{S_m\}$ *be as described in* (h_1, h_2). *If* ρ_m *and* ρ *are the numbers given in* (8) *and* (9), *then* $\lim_{m\to\infty} \rho_m = \rho$.

Proof. Let $P(x, A^*)$ be an element in $\mathbf{P}_k$ such that

$$\| L[P(x, A^*)] \|_S = \inf_{\mathbf{P}_k} \sup_S \mid L[P(x, A)] \mid = \rho^*.$$

Then since S is dense in I,

$$\sup_S \mid L[P(x, A^*)] \mid = \sup_I \mid L[P(x, A^*)] \mid.$$

Thus $\rho^* = \rho$. Now let $x_0 \in I$ be such that

$$\sup_I \mid L[P(x, A_m)] \mid = \mid L[P(x_0, A_m)] \mid, \tag{11}$$

and let $z_m \in S_m$ be such that

$$\mid x_0 - z_m \mid = \min_{s_i \in S_m} \mid x_0 - s_i \mid. \tag{12}$$

Then by (9) and (11)

$$\rho \leqslant \mid L[P(x_0, A_m)] \mid.$$

Let $H(x, y, y') = F(x, y, y') + G(x, y, y')$. Then

$$\begin{aligned} \rho \leqslant{} & \mid h(x_0) - h(z_m) \mid + \mid P''(x_0, A_m) - P''(z_m, A_m) \mid \\ & + \mid H(x_0, P(x_0, A_m), P'(x_0, A_m)) - H(z_m, P(z_m, A_m), P'(z_m, A_m)) \mid \\ & + \mid L[P(z_m, A_m)] \mid. \end{aligned} \tag{13}$$

Because of Lemma 2 we have for $x \in I$ and all m that

$$\mid P(x, A_m) \mid \leqslant N_1, \qquad \mid P'(x, A_m) \mid \leqslant N_2,$$

where N_1 and N_2 are constants. Let

$$\delta_1(m) = \mid H(x_0, P(x_0, A_m), P'(x_0, A_m)) - H(x_0, P(z_m, A_m), P'(z_m, A_m)) \mid$$

and

$$\delta_2(m) = | H(x_0, P(z_m, A_m), P'(z_m, A_m)) - H(z_m, P(z_m, A_m), P'(z_m, A_m))|.$$

Then (13) implies that

$$\rho \leqslant | h(x_0) - h(z_m)| + | P''(x_0, A_m) - P''(z_m, A_m)| + \delta_1(m) + \delta_2(m) + \rho_m . \tag{14}$$

Then the equicontinuity of $\{P(x, A_m)\}$, $\{P'(x, A_m)\}$, the continuity of h, the uniform continuity of H on $I \times [-N_1, N_1] \times [-N_2, N_2]$, (h_2), (12), and (14) imply that

$$\rho \leqslant \lim_{m\to\infty} \rho_m .$$

But for all m,

$$\rho_m \leqslant \rho.$$

Therefore $\lim_{m\to\infty} \rho_m = \rho$.

We conclude this section with the following corollary to Theorem 2.

COROLLARY. *Let* $\{P(x, A_m)\}$ *be a sequence from* $\mathbf{P}_k$ *satisfying* (10) *for each* m. *Then there exists a subsequence* $\{P(x, A_{m_l})\}$ *that converges uniformly on* I *to a* $P(x, A') \in \mathbf{P}_k$. *Furthermore,*

$$\sup_I | L[P(x, A')]| = \rho.$$

The proof follows from (h_1, h_2), Lemma 2, and Theorem 2.

It should be noted that if for a particular operator L the best approximation on S_m to $y(x)$ is unique for all m sufficiently large, and if the best approximation to $y(x)$ on I is unique, then the Corollary implies that $\lim_{m\to\infty} P(x, A_m) = P(x, A)$ uniformly on I, where $P(x, A_m)$ and $P(x, A)$ are the best approximations from $\mathbf{P}_k$ to $y(x)$ on S_m and I, respectively.

5. An Example

The following example illustrates Theorem 2 and the corollary. Let

$$Ly \equiv y'' - (6/(x + 1)^6)\, y^2 = 0, \tag{15}$$

where

$$y(0) = 1, \qquad y'(0) = 3. \tag{16}$$

The solution to (15) and (16) is unique on $I = [0, 1]$. Select $G(x, y, y') = -(6/(x+1)^6)\, y^2$, $F(x, y, y') \equiv 0$, and $h(x) \equiv 0$. Then $u(x) = 6/(x+1)^6$, and $\varnothing(y, y') = y^2$. Hence $1 < \alpha \leqslant 2$, and η is any constant such that $\eta < \alpha$. Let $\mathbf{P}_2 = \{P_2(x, A)\}$, where $A = (1, 3, a_2)$, and where $P_2(x, A) = 1 + 3x + a_2x^2$. Then we wish to best approximate the solution to (15) and (16) in the sense that

$$\| L[P_2(x, A)]\|_I = \sup_I \mid 2a_2 - (6/(x+1)^6)(1 + 3x + a_2x^2)^2 \mid$$

is a minimum over $\mathbf{P}_2$. Theorem 1 guarantees that there exists a $P_2(x, A^*) = 1 + 3x + a_2{}^*x^2$ such that

$$\| L[P_2(x, A^*)]\|_I = \inf_{\mathbf{P}_2} \sup_I \mid L[P_2(x, A)]\| = \rho.$$

Theorem 1 also guarantees that if $\{S_m\}$ is sequence of sets satisfying (h_1, h_2), then for each m there exists a $P_2(x, A_m) = 1 + 3x + a_{2m}x^2$ such that

$$\| L[P_2(x, A_m)]\|_{S_m} = \inf_{\mathbf{P}_2} \sup_{S_m} \mid L[P_2(x, A)]\| = \rho_m .$$

The conclusion of Theorem 2 guarantees that $\lim_{m\to\infty} \rho_m = \rho$, and in this example the Corollary guarantees that

$$\lim_{m\to\infty} \| P_2(x, A_m) - P_2(x, A^*)\|_I = \lim_{m\to\infty} \mid a_2{}^* - a_{2m} \mid = 0.$$

In the following computations all numbers are rounded to three decimal places. Let

$$S_1 = \{0, 0.2, 0.5, 0.8\},$$

$$S_2 = S_1 \cup \{0.1, 0.3, 0.4, 0.6, 0.7, 0.9\},$$

and

$$S_3 = S_2 \cup \{0.05, 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85, 0.95, 1.0\}.$$

Then the best approximation to $y(x)$ on S_1 is

$$P_2(x, A_1) = 1 + 3x + 2.643x^2,$$

and $\rho_1 = 1.149$. The best approximation to $y(x)$ on S_2 is

$$P_2(x, A_2) = 1 + 3x + 2.564x^2,$$

and $\rho_2 = 1.089$. On S_3 the best approximation to $y(x)$ is

$$P_2(x, A_3) = 1 + 3x + 2.486x^2,$$

and $\rho_3 = 1.028$. The best approximation to $y(x)$ on $I = [0, 1]$ is

$$P(x, A^*) = 1 + 3x + 2.486x^2,$$

and $\rho = 1.028$. Thus a_2^* and a_{23} agree to three decimal places, as do ρ and ρ_3.

Acknowledgments

The author is indebted to Professor John A. Brown, Montana State University, for his valuable suggestions regarding certain aspects of this paper, and to the referee, who suggested revisions that have improved the exposition of this paper.

References

1. E. W. Cheney, "Introduction to Approximation Theory," McGraw–Hill, New York, 1966.
2. M. S. Henry, A best approximate solution of certain nonlinear differential equations, *SIAM J. Numer. Anal.* **6** (1969), 143–148.
3. M. S. Henry, Best approximate solutions of nonlinear differential equations, *J. Approximation Theory* **1** (1970), 59–65.
4. R. G. Huffstutler and F. M. Stein, The approximate solution of certain nonlinear differential equations, *Proc. Amer. Math. Soc.* **19** (1968), 998–1002.
5. J. R. Rice, "The Approximation of Functions," Vol. II, Addison–Wesley, Reading, MA, 1969.
6. F. M. Stein and K. F. Klopfenstein, Approximate solutions of a system of differential equations, *J. Approximation Theory* **1** (1968), 279–292.